

UNIVERSITY OF ALBERTA LIBRARY



0 0003 4691 287

# Provincial Report

Grade 3 Social Studies

Achievement Test

October 1984

# Student Evaluation

LB  
3054  
C2  
D3334  
1984

**Alberta**  
EDUCATION

EDUC

DISTRIBUTION:

Superintendents of Schools  
School Principals and Teachers  
The Alberta Teachers' Association  
Alberta School Trustees' Association  
Alberta Education  
General Public Upon Request

Ex LIBRIS  
UNIVERSITATIS  
ALBERTAE



## EXECUTIVE SUMMARY

### Description of the Test

The Grade 3 Social Studies Test was based on the *1981 Alberta Social Studies Curriculum*. All test questions were drawn from the content of the three topics prescribed for Grade 3. These topics are: Interdependence of Communities in Canada Today, Lifestyles of Canadians in Other Times, and Lifestyles in Culturally Distinctive Communities.

The test consisted of 50 multiple-choice questions, which were answered directly in the test booklet. There was no separate answer sheet. The test was organized into three approximately equal sections, each section related to one of the prescribed topics.

### Administration

The test was administered on June 12, 1984 to 27 789 students. A small percentage of the student population in split-grade classes wrote only the part of the test that they had been taught.

### Results

The provincial average is 31.4 out of 50 or 62.8%. The standard deviation is 8.4.

Although overall student achievement in Grade 3 Social Studies at the provincial level was acceptable, the level of achievement was not uniform among reporting categories. Student achievement was lowest on Topic C, the section that assessed knowledge and skills related to the study of culturally distinctive communities.

In a typical Grade 3 classroom, 85% of the students should be able to achieve the minimum objectives of the program. On this test 44% is considered to be the minimum acceptable level of achievement in meeting course objectives. Over 85% of the students achieved at or above this minimal level.

## TABLE OF CONTENTS

	PAGE
ACKNOWLEDGMENTS	vi
CHAPTER 1: THE ACHIEVEMENT TESTING PROGRAM	1
Exemptions from the Achievement Testing Program	1
CHAPTER 2: TEST DESIGN, DEVELOPMENT, AND DESCRIPTION	2
Test Design and Development Process	2
Test Description	3
Approach to Optional Content	3
Reporting Categories	5
CHAPTER 3: ADMINISTRATION OF THE TEST	7
Determination of the Student Population	7
Administration	7
Data Collection	7
Standard-Setting	8
CHAPTER 4: RESULTS AND OBSERVATIONS	9
Test Results	9
Comparison of Results to Standards	10
Results for Students Completing Only Portions of the Program	11
Total Score Frequency Distribution	11
Response Frequencies for Individual Questions	13
Discussion of Selected Questions	14
CHAPTER 5: GUIDE TO THE INTERPRETATION OF JURISDICTION RESULTS	27
Differences Between Jurisdiction and Provincial Averages	27
Absentee Rates	30

LIST OF TABLES

TABLE		PAGE
1	Number and Distribution of Questions	3
2	Test Blueprint	4
3	Reporting Categories -- Question Distribution	6
4	Results by Reporting Category	9
5	Comparison of Results to Standards	10
6	Comparison of Results -- Total and Partial Completion of the Program	11
7	Total Score Frequency Distribution	12
8	Response Frequencies	13
9	Distribution of Jurisdiction Levels of Achievement	29

#### ACKNOWLEDGMENTS

The successful administration of the Grade 3 Social Studies Achievement Test was due to the concerted effort of all involved. Success would have been impossible without substantial contributions from many people, particularly the administrators, teachers, and students, who extended their full co-operation.

The advice received from the Test Review Committee regarding design, development, and reporting has been particularly valuable in the implementation of the Achievement Testing Program. This Committee has representation from:

The Alberta Teachers' Association  
The Conference of Alberta School Superintendents  
The Universities  
Alberta Education

The contribution made by this group is gratefully acknowledged.

The technical expertise provided by Dr. T. O. Maguire, Professor, Division of Educational Research Services, University of Alberta, has also contributed greatly to the advancement of the Achievement Testing Program, and his work in this area is acknowledged and appreciated.

Lloyd E. Symyrozum  
Director  
Student Evaluation Branch

## Chapter 1

### THE ACHIEVEMENT TESTING PROGRAM

The purpose of the Achievement Testing Program is to provide educators, trustees, and others with information, significant at the provincial and local levels, about student knowledge, understanding, and skills in relation to program objectives.

The achievement tests are specific to the program of studies prescribed by the Minister of Education. Curriculum specifications for each subject area, provided by the Curriculum Branch and the Language Services Branch of Alberta Education, identify the major content areas, the specific learning objectives within each area, and the emphasis that each objective is to receive. The test questions reflect these curriculum specifications.

The achievement tests are administered on a cyclical basis in four subject areas: language arts, social studies, mathematics, and science, and at three grade levels: 3, 6, and 9. In 1984, achievement tests were administered in Grade 3 Social Studies, Grade 6 English Language Arts, and Grade 9 Mathematics.

Following the achievement test administration in June of each year, the results are reported to each school jurisdiction. These district profiles include results for each school and each student, but individual statements of results are not issued to students.

This report is designed to assist school jurisdictions in interpreting their results.

#### Exemptions from the Achievement Testing Program

Under normal circumstances, the following are exempt from achievement testing:

- Students for whom grants are received from the Special Educational Services Branch
- Students in classes in which the subject being tested has been cycled and taught in an alternate year
- Students in classes in which the subject being tested has been taught in an alternate semester
- Students enrolled in English as a Second Language programs for whom grants are received under Section 54(2) of the *School Grants Regulations*

## Chapter 2

### TEST DESIGN, DEVELOPMENT, AND DESCRIPTION

#### Test Design and Development Process

The development of the Grade 3 Social Studies Achievement Test progressed through four stages: preparation of curriculum specifications, design and development of the test, construction and administration of the pilot test, and construction of the final test form.

##### 1. Curriculum Specifications

The Curriculum Branch of Alberta Education prepared curriculum specifications that identify the major content areas, the specific objectives within each area, and the emphasis each is to receive in the classroom. The curriculum specifications were distributed to all school jurisdictions in the province in the publication *Grade 3 Social Studies Curriculum Specifications*.

##### 2. Test Design and Development

The Student Evaluation Branch of Alberta Education developed a test blueprint and reporting categories based on the weightings of the prepared curriculum specifications. Test questions were developed by Grade 3 Social Studies teachers from all parts of the province under the supervision of the Student Evaluation Branch. Revisions were made to the questions on the basis of teacher recommendations and field test results.

##### 3. Pilot Test

A pilot test was constructed from the bank of field-tested questions and administered to approximately 250 Grade 3 students in different parts of the province. Revisions were made to the test on the basis of the pilot test results by the Student Evaluation Branch, in consultation with Social Studies supervisors, consultants, and classroom teachers. The final test design, complete with blueprint and sample questions, was presented to all school jurisdictions in the province in the publication *Student Achievement Testing Program: Grade 3 Social Studies* (Student Evaluation Branch Bulletin, Volume 3, Number 9, November 1983).

##### 4. Final Test Form

A draft of the final test was constructed from those questions that best reflected curricular intent and test design requirements. The Test Review Committee reviewed the draft for content validity, accuracy and technical merit. Final changes were made to the test, taking into consideration the recommendations of this committee.

### Test Description

The test was based on the 1981 Alberta Social Studies Curriculum for Grade 3. Content emphasis was derived from the Grade 3 Social Studies Curriculum Specifications, 1981.

The test consisted of 50 multiple-choice questions, which were answered in the test booklet. There was no separate answer sheet.

The topics and objectives on which the questions on the test were based are shown in the table below. (See also the test blueprint on page 4.)

Table 1

#### Number and Distribution of Questions Grade 3 Social Studies Achievement Test

	<u>Value</u>	<u>Knowledge</u>	<u>Skill</u>	<u>Total</u>	<u>Emphasis in %</u>
Topic A	2	7	8	17	34
Topic B	2	8	7	17	34
Topic C	2	7	7	16	32
<b>TOTAL</b>	<b>6</b>	<b>22</b>	<b>22</b>	<b>50</b>	
<b>Emphasis in %</b>	<b>12</b>	<b>44</b>	<b>44</b>		<b>100%</b>

An attempt was made to minimize the use of difficult vocabulary. Readability was carefully monitored, both through the involvement of numerous Grade 3 classroom teachers during the test development process and through the application of the Spache Readability Formula. Allowances were made for vocabulary specific to the Alberta Grade 3 Social Studies program.

### Approach to Optional Content

The curriculum allows for the selection of a variety of communities for study in Topics A and C. Since many different communities were studied in Grade 3 classrooms throughout the province, it was not possible to test students' recall of specific facts about particular communities. Students were therefore tested on their recall and comprehension of general information related to the prescribed concepts and generalizations. To do well, students had to understand that the communities they studied were only examples of many similar communities. For example, the culturally distinctive communities about which they had learned have many of the same characteristics as other such communities.

To avoid the possibility of over-generalizing about the lifestyles of individuals or groups within any one culturally distinctive community, some questions in Topic C were related to a short reading about an imaginary community.

Table 2  
Test Blueprint\*

#### CONCEPT REPORTING CATEGORY

#### 4. B. VALUE CONCEPTS:

## Reporting Categories

To provide information that is both meaningful and reliable, it is necessary to group questions into reporting categories (subtests). Each reporting category requires a minimum of six questions to achieve statistical reliability. The reporting categories used in the Grade 3 Social Studies Achievement Test are described below. The distribution of questions by category is presented in Table 3 on page 6.

1. TOPIC A: All questions related to the interdependence of communities in Canada today.
2. TOPIC B: All questions related to lifestyles of Canadians in other times.
3. TOPIC C: All questions related to lifestyles of Canadians in culturally distinctive communities.
4. RECALL AND COMPREHENSION (ALL TOPICS): Recalls and understands facts, concepts, and generalizations.
5. RECALL AND COMPREHENSION (TOPIC A): Recalls and understands facts, concepts, and generalizations about the interdependence of communities in Canada today.
6. RECALL AND COMPREHENSION (TOPIC B): Recalls and understands facts, concepts, and generalizations about lifestyles of Canadians in other times.
7. RECALL AND COMPREHENSION (TOPIC C): Recalls and understands facts, concepts, and generalizations about lifestyles of Canadians in culturally distinctive communities.
8. VALUE CONCEPTS AND VALUING SKILLS (ALL TOPICS): Recalls and understands competing values and analyses competing value positions.
9. INQUIRY SKILLS I (ALL TOPICS): Uses skills related to the early stages of the inquiry process (i.e., identifying an issue, gathering and organizing data).
10. INQUIRY SKILLS II (ALL TOPICS): Uses skills related to the later stages of the inquiry process (i.e., analysing data, resolving an issue).

Table 3

## Distribution of Questions by Reporting Category\*

Category	Questions	Number of Questions
Topic A	1 - 17	17
Topic B	18 - 34	17
Topic C	35 - 50	16
Recall and Comprehension (all topics)	6,7,8,9,10,12, 13,20,21,22,23, 24,25,26,27,35, 36,37,39,45,46,48	22
Recall and Comprehension (Topic A)	6,7,8,9,10,12,13	7
Recall and Comprehension (Topic B)	20,21,22,23,24,25, 26,27	8
Recall and Comprehension (Topic C)	35,36,37,39,45, 46,48	7
Value Concepts and Valuing Skills (all topics)	3,11,28,29,38,47	6
Inquiry Skills I (all topics)	1,2,5,14,15,18, 30,32,33,41,42, 43,49	13
Inquiry Skills II (all topics)	4,16,17,19,31,34, 40,44,50	9

\* Some questions are reported in more than one category.

## Chapter 3

### ADMINISTRATION OF THE TEST

#### Determination of the Student Population

The larger school jurisdictions could choose to test either all Grade 3 Social Studies students or only randomly selected schools. School boards wanting to sample student achievement were required to notify the Student Evaluation Branch. Only one jurisdiction opted for sampling.

Some Grade 3 students (students in their third year, taking the Grade 3 program) in split-grade classes did not receive instruction in all three social studies topics because the program was cycled with another grade over a two-year period. These students were exempt from writing those topics on the test that they had not been taught, provided that the exemptions had been approved by the superintendent.

#### Administration

Jurisdictions were requested in April 1984 to report the number of students enrolled in Grade 3 Social Studies in each school. In May, letters were sent by the Student Evaluation Branch to superintendents, principals, and teachers in the province requesting their co-operation in the testing. Information addressed to the superintendents and principals included the test schedule, procedures for test administration, and requirements for returning test materials. Information addressed to the teachers related to the administration of the test and the return of test materials. Each jurisdiction was sent the appropriate number of tests and administration instructions, packaged according to school. After the test was administered, teachers were instructed to collect all test booklets and return them to the principal for forwarding to school board offices, which, in turn, were responsible for sending the test booklets to the Student Evaluation Branch.

Staff from the Regional Offices of Education supervised the administration of the test in private schools.

#### Data Collection

A total of 920 schools from 142 public and separate school jurisdictions returned scorable booklets for 27 118 students. A total of 71 schools from 69 private jurisdictions returned scorable booklets for 671 students.\*

---

\*There were 2001 students from public and separate jurisdictions and 13 students from private jurisdictions who only completed portions of the program, and therefore only wrote parts of the test.

### Standard-Setting

While provincial averages are useful for comparing the scores of students in a particular school or jurisdiction with overall levels of achievement, it is not possible to know whether the students in the province did as well as they should. A test score by itself has limited meaning without comparison to a standard. Tests vary in difficulty: a raw score of 25/50 for example, could represent very high achievement on one test, and very low achievement on another.

To establish a standard that allows the assessment of overall achievement on the test, the Student Evaluation Branch follows certain procedures. For the Grade 3 Social Studies test, experienced Grade 3 teachers from all parts of the province met to determine what raw score would be expected on the test for a borderline student. The borderline is the division between those who could be expected to achieve the minimum objectives, and those who could not. After a review of the curriculum, it was judged that 85% of Grade 3 students should be able to achieve the minimum goals of the Grade 3 Social Studies curriculum, given adequate teaching and resources. Since 85% of students should be able to reach this level, the borderline student would be at the 15th percentile in ability.

The teachers examined each question on the test and determined the difficulty of that question for a 15th percentile student. From the individual question difficulties, the overall test difficulty for the borderline student was determined. The average of the test difficulties established by the teachers is the standard for the test. For the Grade 3 Social Studies Achievement Test the standard established was as follows: Given the nature and difficulty of this test, 85% of the students should achieve a score of 44% or better.

## Chapter 4

### RESULTS AND OBSERVATIONS

#### Test Results

The results of the Grade 3 Social Studies Achievement Test are summarized by reporting category in Table 4 below. These figures present only the results for those students writing the total test.

Table 4  
Results by Reporting Category

Reporting Category	No. of Questions	Mean	Standard Deviation
1. Total Test Score	50	31.4	8.4
2. Topic A: Interdependence of Communities in Canada Today	17	10.4	3.0
3. Topic B: Lifestyles of Canadians in Other Times	17	11.5	3.2
4. Topic C: Lifestyles in Culturally Distinctive Communities	16	9.4	3.4
5. Recall and Comprehension (all topics)	22	14.1	3.9
Recall and Comprehension (Topic A)	7	4.1	1.4
Recall and Comprehension (Topic B)	8	5.8	1.7
Recall and Comprehension (Topic C)	7	4.1	1.8
6. Value Concepts and Valuing Skills (all topics)	6	3.9	1.4
7. Inquiry Skills I (all topics)	13	7.8	2.5
8. Inquiry Skills II (all topics)	9	5.6	2.2

Although performance on different reporting categories of the test shows some variation, these scores are not directly comparable. The sets of questions that make up the different categories were not selected to be equal in average level of difficulty, therefore differences may be due to variations in question difficulty rather than in student performance. In combination with jurisdictional results, however, the norms can be used to detect patterns of relative strength or weakness in student achievement.

#### Comparison of Results to Standards

The minimum acceptable standard (the score representing borderline achievement) set for the Grade 3 Social Studies Achievement Test by teachers was a raw score of 22 out of 50. Thus, 85% of students should be at or above this standard, given adequate teaching and resources.

Table 5 presents the percentage of students scoring at or above the minimum acceptable standard in a number of reporting categories. These figures report only the results for those students who wrote the total test.

Table 5

#### Comparison of Results to Standards

Reporting Category	Raw Score Representing Borderline Achievement	% of Students Scoring at or above Borderline Level
Total Test	22/50	85.7
Recall and Comprehension (all topics)	10/22	86.1
Value and Skill Objectives	13/28	81.2
Topic A	8/17	81.9
Topic B	8/17	87.9
Topic C	8/16	70.1

The number of students meeting the standard was slightly greater than the expected proportion of 85% on the total test and on two of the major reporting categories: Topic B, and Recall and Comprehension. The proportion of students meeting the standard was slightly less than expected on Topic A and on Value and Skill Objectives, and significantly less on Topic C.

### Results for Students Completing Only Portions of the Program

Table 6 presents and compares by topic the results for students who wrote only portions of the test with the results for students who completed the total program and therefore wrote the entire test.\*

Table 6

#### Comparison of Results -- Total and Partial Completion of the Program

	Students Completing the Total Program			Students Completing Only Portions of the Total Program		
	No. of Students	Mean	Standard Deviation	No. of Students	Mean	Standard Deviation
Topic A	25 775	10.5	3.0	958	10.2	3.4
Topic B	25 775	11.6	3.2	1169	11.3	3.5
Topic C	25 775	9.5	3.4	805	9.6	3.2

There was only a slight difference in the level of achievement between the two groups on all three topics.

### Total Score Frequency Distribution

Table 7, on the following page, presents the frequency distribution of scores for those students writing the total test. Any score that was achieved by fewer than .05% of the population is considered to be a relative frequency of 0.0, and is represented by a dash. Therefore it should be noted that the range of student scores was from 2 to 50, although the relative frequencies at the top and lower ends of the distribution do not appear to indicate this. For example, eight students achieved a raw score of 50, but since this represents fewer than .05% of the population, the relative frequency is considered to be 0.0 and is indicated by a dash.

---

\*Students in some split-grade classes did not receive instruction in all three topics. These students were exempt from writing those topics on the test that they had not been taught.

Table 7  
Total Score Frequency Distribution

Score	Relative Frequency in %*	Cumulative Frequency in %**	Score	Relative Frequency in %*	Cumulative Frequency in %**
2	-	-	27	3.4	31.5
3	-	-	28	3.5	35.0
4	-	-	29	4.0	39.0
5	-	0.1	30	4.2	43.2
6	-	0.1	31	3.9	47.1
7	0.1	0.2	32	4.4	51.5
8	0.1	0.3	33	4.4	55.9
9	0.1	0.4	34	4.6	60.5
10	0.2	0.6	35	4.3	64.8
11	0.3	0.9	36	4.5	69.3
12	0.5	1.4	37	4.3	73.6
13	0.7	2.1	38	4.1	77.7
14	0.9	3.0	39	4.1	81.8
15	1.0	4.0	40	3.5	85.3
16	1.1	5.1	41	3.2	88.5
17	1.5	6.6	42	3.0	91.5
18	1.5	8.1	43	2.5	94.1
19	1.8	9.8	44	2.0	96.1
20	2.0	11.8	45	1.6	97.7
21	2.4	14.3	46	1.1	98.8
22	2.5	16.7	47	0.7	99.5
23	2.6	19.3	48	0.3	99.8
24	2.8	22.1	49	0.1	99.9
25	2.9	25.0	50	-	100.0
		TOTAL		100.0	

\*Relative Frequency: the percentage of students who obtained each score.

\*\*Cumulative Frequency: the percentage of students scoring at or below each score.

Response Frequencies for Individual Questions

The response frequencies for all 50 questions appearing in the test are presented in Table 8. These figures present only the response frequencies for those students writing the total test.

Table 8  
Response Frequencies

Question Number	Key	Distribution of Responses in %*				Question Number	Key	Distribution of Responses in %*			
		A	B	C	D			A	B	C	D
1	B	23	53	22	2	26	A	74	11	10	3
2	C	7	27	46	20	27	D	3	14	6	74
3	B	16	59	2	21	28	B	3	77	2	16
4	A	77	5	10	6	29	C	15	4	76	3
5	C	5	18	70	6	30	C	10	16	62	10
6	D	10	8	1	81	31	A	44	9	29	15
7	C	40	7	20	32	32	D	12	12	29	46
8	D	3	3	2	92	33	A	78	7	3	10
9	A	49	8	23	16	34	B	3	76	10	7
10	D	33	4	6	56	35	C	8	30	56	6
11	B	8	65	12	14	36	D	7	12	11	68
12	B	34	49	5	11	37	B	12	71	9	6
13	C	14	6	59	20	38	B	13	77	5	4
14	A	65	11	15	7	39	A	52	9	18	19
15	A	75	4	2	9	40	D	20	14	10	55
16	A	69	8	15	6	41	D	2	44	4	49
17	D	5	18	16	57	42	A	57	21	17	5
18	C	7	5	65	17	43	C	18	19	52	8
19	D	12	11	22	53	44	A	69	13	9	7
20	B	13	77	3	6	45	C	15	10	62	11
21	A	76	4	1	12	46	B	18	67	8	6
22	D	6	12	10	71	47	D	9	23	33	34
23	C	6	6	73	14	48	C	6	34	51	7
24	A	71	10	7	11	49	D	5	30	3	60
25	B	12	60	17	10	50	A	64	12	8	12

\*The sum of the percentages may be less than 100% because the No Response category is not included and the numbers have been rounded.

## Discussion of Selected Questions

Selected questions from the Grade 3 Social Studies Achievement Test follow. They illustrate the nature and complexity of questions within the major reporting categories. The darkened circle in front of or below an alternative indicates the keyed response for each question. The percentage of students who selected each alternative is given in brackets.

1. RECALL AND COMPREHENSION: (Questions: 6,7,8,9,10,12,13,20,21,22,23,24, 25,26,27,35,36,37,39,45,46,48)

Questions in this reporting category measure a student's ability to recall and understand:

- examples of the types of services found in communities
- characteristics of specialization
- characteristics of interdependence among communities
- similarities and differences between urban and rural communities
- examples of changes that occur as communities become urbanized
- how and why people immigrated to Alberta
- where early settlements were located
- examples of the goals and aspirations of early settlers
- examples of how early settlers met their needs
- changes to the lifestyles of people who immigrated
- how settlers worked together to build a community
- features of culturally distinctive communities and how they are preserved
- differences and similarities among different cultures in Alberta
- effects on culturally distinctive communities of interaction with other cultures

Student achievement in this reporting category was slightly higher than the average for the total test. Differences in levels of achievement occurred among the three topics. The average score for the questions on Topic B was higher than the average scores for Topic A or Topic C, and the range of the difficulty of items was smaller in Topic B than in either Topic A or Topic C.

The curriculum allows for the study of a variety of communities in topics A and C. For this reason, the questions from these topics tested students' recall and comprehension of the prescribed concepts and generalizations in a less specific way than did the questions from Topic B. This discrepancy may account, in part, for the differences in level of student achievement that occurred among the three topics.

Several questions from this reporting category are presented and discussed on the following pages.

Question 7 tested students' understanding of an attribute of the concept specialization (Topic A). It appeared on the test as follows:\*

7. Mike made the list below in his notebook. Which city in this list is the **MOST** specialized?

(40%)  Edmonton — oil  wheat  and beef 

(7%)  Montreal — shipping  and clothing 

(20%)  Oshawa — cars 

(32%)  Vancouver — shipping  fish  and lumber 

This question proved to be very difficult for the students. More students chose alternatives A and D than the keyed response, C. However, C is clearly the correct answer. Students may have selected alternatives A and D in response to the word "MOST," which appeared in the stem. It appears that they selected a city having the greatest number of industries rather than the one that was most specialized. This would seem to indicate that most students were unaware that an essential attribute of a specialized community is that fewer types of industries are present.

---

\*This question and others from the achievement test that appear in this report have been reduced in size. They appear much larger in the actual test booklet.

Question 20 is an example of another item in this reporting category. It tested students' recall of a historical fact about early settlers (Topic B). It appeared on the test as follows:

20. Most early settlers came to Alberta to



(13%)  drill for oil



(77%)  farm the land



(3%)  look for gold



(6%)  start factories

An analysis of the response patterns shows that students representing a wide range of achievement levels were able to answer this question correctly. This suggests that most students were taught the specific fact being tested.

Question 37 tested students' ability to recall the role of special holidays in maintaining traditions in culturally distinctive communities (Topic C). It appeared on the test as follows:

37. Why do some communities have their own special holidays?

- (12%)  To show that their way of life is more fun than that of other people
- (71%)  To remember important events in their history or religion
- (9%)  To have more days to give and receive gifts
- (6%)  To have more holidays than other people

Most students appeared to understand the importance of special holidays in maintaining traditions.

2. VALUE CONCEPTS and VALUING SKILLS: (Questions: 3,11,28,29,38,47)

Questions in this reporting category measure a student's ability to:

- identify an example of self-reliant behavior
- recall reasons for valuing self-reliance
- select a historical example of co-operative behavior
- identify co-operation as the value underlying a behavior
- identify tolerance and a willingness to learn about others as being values of importance to some people
- identify a behavior that indicates a desire to maintain a culturally distinctive lifestyle

Student achievement in this reporting category was slightly above the average score for the total test. The difficulty level of the questions in this category ranged from .59 to .77 with the exception of question 47, which proved to be considerably more difficult for students. This item is presented and discussed below.

Question 47 was based on a story about Lundi, a mythical culturally distinctive community. Students were required to identify a behavior that best indicated that the Lundians wanted to maintain their culturally distinctive lifestyle. The question appeared on the test as follows:

**47. What tells us that the Lundians want to keep their community special?**

(9%)  **Each family has its own house.**  
(23%)  **They do not watch television.**  
(33%)  **Each family grows all of its own food.**  
(34%)  **They send their children to school on Saturday.**

The statistics show that many students were attracted to alternatives B and C. These alternatives described characteristics of the Lundians, but ones that were less directly related to the preservation of a culturally distinctive lifestyle than was the keyed answer D, which implied the study of the customs and language unique to their culture. Students who chose alternative B or C may have had difficulty making this distinction.

Question 11 required students to identify a reason for choosing to use only the goods produced in one's own community. It appeared on the test as follows:

11. Why would some people choose to use only the goods made in their own community?

(8%)  To have less work to do  
(65%)  To depend more on themselves  
(12%)  To use up the extra goods made  
(14%)  To have many kinds of goods

This item was less difficult for students than was question 47.

3. INQUIRY SKILLS I: (Questions: 1,2,5,14,15,18,30,32,33,41,42,43,49)

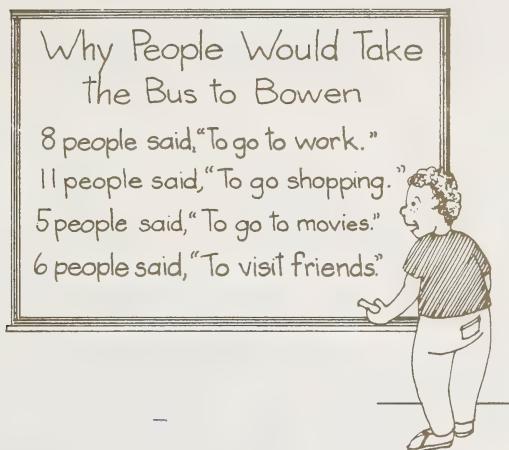
Questions in this reporting category tested a student's ability to:

- determine distances using non-standard measures
- identify appropriate research materials
- translate data from written to graphic form
- record events on a timeline
- read and interpret a chart
- record information on a chart
- paraphrase a statement of a problem
- describe the location of a community using cardinal directions

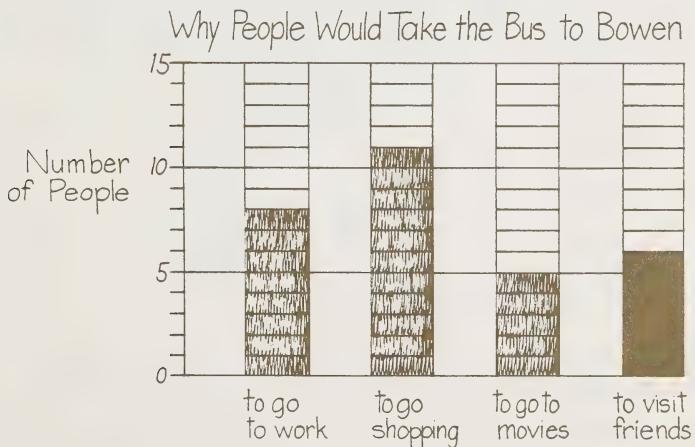
The average score for the questions in this category was slightly lower than the average for the total test. The lowest scores were on some of the questions that dealt with maps and graphs. Two questions from this reporting category are presented and discussed on the following pages.

Question 15 tested students' ability to represent data graphically. Students had to translate data from written form into graphic form. This question appeared as part of a series of questions related to the same issue. It appeared on the test as follows on page 19.

15. Mike asked 30 people why they want the new bus service. He made a list for the class to show what he found.



Mike wanted to make a graph to show what he found. Finish the bar graph below for Mike.

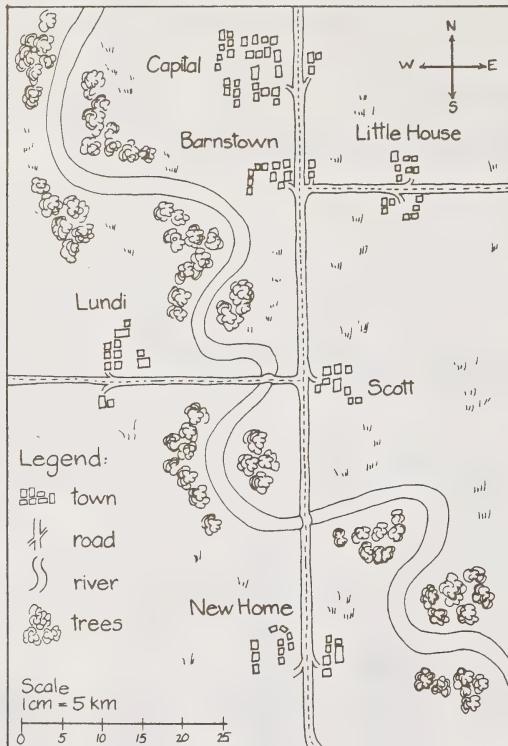


Students had to darken the spaces on a bar rather than fill in a circle to indicate their responses to this item. Seventy-five per cent of the students completed the graph in a satisfactory way (i.e., darkened the bottom 6 spaces). Four per cent filled in 5 spaces or fewer and two per cent filled in 7 spaces or more. Nine per cent completed the graph in an irregular fashion, and eleven per cent did not respond at all.

Question 43 was one of a number of questions testing students' ability to read and interpret maps using non-standard measure. The map and question appeared on the test as follows:

Imagine that there is a special community in Alberta called Lundi. Most of the people of Lundi came from the same village in Europe. They call themselves "Lundians."

Use the map below to answer questions 41 to 43.



43. Find Lundi on the map. Which town is 25 km north and 40 km east of Lundi?

- (18%)  Barnstown
- (19%)  Capital
- (52%)  Little House
- (8%)  New Home

A number of teachers expressed concern about this and similar items appearing on the test because of the scales shown on the maps. They questioned the validity of these questions since the curriculum does not specifically mention the use of scale at the Grade 3 level.

The distribution of student responses indicates that question 43 fell well within the acceptable limits of difficulty. Further analysis revealed that this question discriminated well between the upper and lower groups on the test, as did the other map questions. Furthermore, the prescribed text for Grade 3, *Cities Are For People*, uses a similar scale on all of its maps. As a result, it was concluded that this was a valid and reliable question for measuring students' ability to use non-standard measures and therefore should be retained. All other questions on using scale were retained for similar reasons.

4. INQUIRY SKILLS II: (Questions: 4,16,17,19,31,34,40,44,50)

Questions in this reporting category tested a student's ability to:

- formulate a generalization about why people are interdependent
- identify and compare points of view on the desirability of increased interdependence
- relate causes and effects of community change over time
- draw conclusions about why lifestyles have changed since 1905
- identify the inherent value underlying a particular course of action
- predict the consequences of learning about other's customs
- summarize information into main points
- formulate a generalization about why people resist change

The average score for the questions in this category was equal to the average score for the total test. Question 31 proved to be the most difficult one in this reporting category for students. It required the students to use information provided in a chart. This question is presented and discussed on page 26 as part of a family of items that appeared on the test.

Question 40 was also part of the Inquiry Skills II reporting category. It was part of a family of questions and tested students' ability to predict a likely consequence of an increase in interaction between two communities. It appeared on the test as follows:

40. What would happen if many people in Wendy's community learned the songs of the special community?

(20%)  Wendy's community would become a special community.

(14%)  Wendy's community would forget all its own songs.

(10%)  The special community would stop singing its old songs.

(55%)  Wendy's community would understand the special community better.

This question was also more difficult than average for students. Students who chose alternative A may have had difficulty understanding the essential nature of a special community.

## 5. FAMILIES OF QUESTIONS

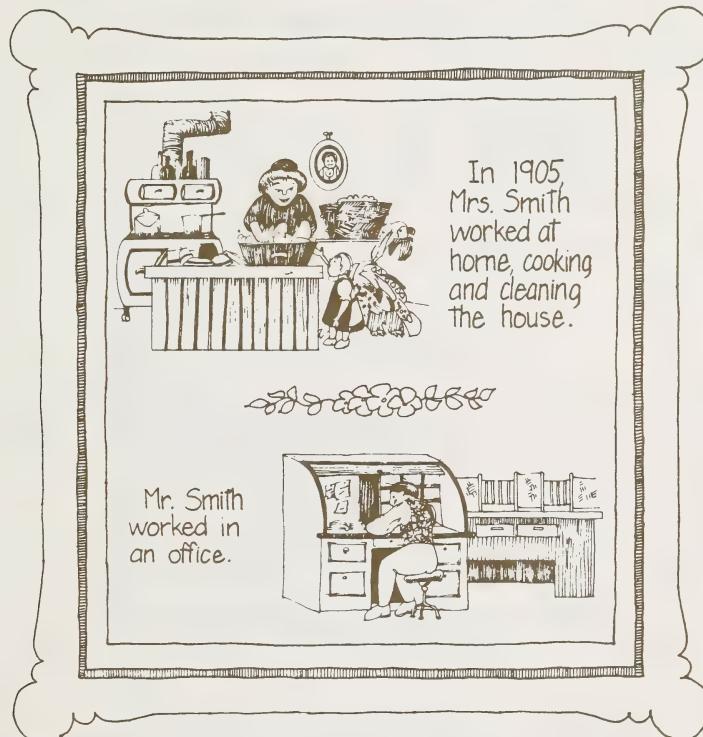
In an attempt to simulate the inquiry process, the test was designed to include series or "families" of questions related to a single issue. Within each family of questions -- and as much as possible throughout each section of the test -- students followed a logical progression from one question to the next similar to the sequence followed in the inquiry process.

Questions 28 to 31, which appear on the following pages, illustrate a typical family of questions from the test. All questions in this family related to the same issue and were based on the same set of characters. Three of the four major reporting categories were represented in this series of questions: Value Concepts and Valuing Skills, Inquiry Skills I, and Inquiry Skills II.

Questions 28 and 29 were part of the reporting category Value Concepts and Valuing Skills.

Question 28 required students to apply their understanding of co-operative behavior by selecting the appropriate example in a new situation (i.e., as described in the historical account). It appeared on the test as follows:

**Use the story below to answer question 28.**



28. The Smiths helped each other when Mrs. Smith took care of their home and Mr. Smith

- (3%)  ate dinner at home
- (77%)  earned money for the family
- (2%)  visited with the neighbors
- (16%)  cleaned up his office

Most students were able to select the correct alternative. The majority of students choosing an incorrect alternative were drawn to the fourth alternative, "cleaned up his office," no doubt in response to the illustration presented with the question.

Question 29 tested students' ability to recognize the underlying value motivating a specific action. In this case, the action is related to the value concept of co-operation.

**Use the story below to answer question 29.**



29. One night Mrs. Brown asked for help with the cleaning because she had to go out. Although Mr. Brown was tired and had his own chores to do, he decided to help.

This shows that it was important for him to

- (15%)  look after his own needs
- (4%)  do things his own way
- (76%)  do something for someone else
- (3%)  make rules for someone else

Students did as well on this question as on question 28.

Questions 30 and 31 were based on the following chart:

	Then (1905)		Now (1984)		
	Mrs. Smith	Mr. Smith	Mrs. Brown	Mr. Brown	
Housework		✓		✓	✓
Office Work			✓	✓	✓

Question 30, which formed part of the reporting category Inquiry Skills I, tested students' ability to read and interpret a chart. It appeared on the test as follows:

**30. What does the chart show?**

- (10%)  Families did more work in 1905 than they do now.
- (16%)  Jobs were easier in 1905 than they are now.
- (62%)  The Browns share more of the same jobs than did the Smiths.
- (10%)  The Smiths worked harder at their jobs than do the Browns.

On this item, the percentages of students who responded incorrectly were distributed more or less equally among the incorrect alternatives. This suggests the absence of a common source of weakness among students who failed to respond correctly.

Question 31 was part of the reporting category Inquiry Skills II. It tested students' ability to draw the best conclusion about why differences exist between the way people lived in 1905 and the way they live now. It appeared on the test as follows:

31. In 1905 most families divided up their work like the Smiths. Today many families divide their work like the Browns.

What might **BEST** explain this?

- (44%)  People have changed their ideas about how to share work.
- (9%)  People want cleaner houses than they did in the past.
- (29%)  People are more willing to work hard today.
- (15%)  People have bigger houses to clean today.

This question was more difficult than most for students. Some students appeared to have difficulty deciding what was a plausible reason for the differences in lifestyles that were presented in the chart.

## Chapter 5

### GUIDE TO THE INTERPRETATION OF JURISDICTION RESULTS

In addition to their use in monitoring student achievement for the province as a whole, the results of the Grade 3 Social Studies Achievement Test are useful in comparing achievement in a particular jurisdiction with provincial results. However, care must be exercised in making these comparisons and in drawing conclusions from the data.

The following jurisdiction and school reports are provided for each jurisdiction under separate cover.

1. The Jurisdiction Summary Report contains jurisdiction equivalents of the provincial results in tables 4, 5, and 7. The jurisdiction equivalent of Table 6 is included for those jurisdictions with classes which did not complete all portions of the program.
2. School Summary Reports contain the school equivalents of the provincial results in tables 4, 5, and 7. The school equivalent of Table 6 is included for those schools with classes which did not complete all portions of the program.
3. The Jurisdiction Item Alternative Response Frequency Data is equivalent to the provincial results in Table 8.
4. The School Item Alternative Response Frequency Data is equivalent to the provincial results in Table 8.
5. Individual Student Sub-scale Results

These reports are confidential to the jurisdiction.

#### Differences Between Jurisdiction and Provincial Averages

Jurisdictions are provided with their average scores for each reporting category. These scores may be compared to the provincial average for the same reporting category. However, the importance of differences between group jurisdiction averages and provincial averages is not always obvious. To aid in the interpretation of differences between the averages, jurisdiction reports indicate when the difference is unlikely to be due to chance variation in the abilities of students. For the purposes of the provincial testing program, the 95% confidence interval is used. That is to say, if the probability is only one in 20 that the difference is due to chance, the jurisdiction average is considered different from the provincial average. Otherwise, it is classified as not different from the provincial average.

A statistical test of significance is made for each reporting category for each jurisdiction. The provincial average for that reporting category and the provincial standard deviation for that category determine the true population average and standard deviation. The standard deviation of a distribution is a measure of the variation of scores. In a normal distribution, there is a fixed and known relationship between the standard deviation and the proportion

of individual scores in any part of the distribution. For example, 68% of scores fall within one standard deviation of the mean (average). If a test has a mean of 50 and a standard deviation of 10, 68% of those writing the test scored between 40 and 60.

The amount of chance variation in jurisdiction averages varies with the number of students tested in that jurisdiction. When any random sample is drawn from a population, its average is expected to be the same as the population average. Yet the actual group average may vary because of individual variation in the sample. This variation is known as the error of the mean. The amount of variation in the averages of random samples drawn from the population is related to the standard deviation of the scores in the population. When the population mean and standard deviation are known, as in the case of the achievement tests, it is possible to determine how likely it is that any subgroup, such as a jurisdiction, represents a random sample of the population in achievement. This statistical test, known as a one-sample z-test, is the one that has been applied to jurisdiction scores in each reporting category. Thus if a jurisdiction is classified as different from the provincial average, there is less than one chance in 20 that the difference between the average score for the jurisdiction on that reporting category and the provincial average would occur in a group of that size selected at random from all students in the province. In other words, the difference is statistically significant at the 0.05 level.

Because these achievement levels are calculated taking jurisdiction size into consideration, two jurisdictions with the same averages but of different sizes may be classified differently. The larger jurisdiction would be more likely to be above or below average, because the amount of chance variation would be less in larger jurisdictions, and the actual difference would represent a larger variation from the provincial average.

For example, imagine two jurisdictions, A with 25 students writing Test X, and B with 100 students writing Test X. Both jurisdictions have the same average, 54.2. Test X has a provincial average score of 50.0 and a standard deviation of 12.0. The difference between the provincial average and the jurisdiction average is 4.2. A difference this large would be expected 8 times out of 100 for groups of 25 selected at random from the population, and fewer than 3 times out of 1000 for groups of 100. Thus the difference from the provincial average would not be statistically significant for Jurisdiction A, but would be for Jurisdiction B.

When it has been determined that a difference is significant, the direction of the difference is important, particularly for those jurisdictions below the provincial average. These jurisdictions are encouraged to identify the sources of these differences.

Table 9 on page 29 indicates the percentage of jurisdictions classified as significantly above or below the provincial average for each reporting category.

Table 9

## Distribution of Jurisdiction Levels of Achievement

Reporting Category	% Below the Provincial Average	% Not Different from Provincial Average	% Above the Provincial Average
Total Test	21.1	50.7	28.2
Topic A	19.0	54.2	26.8
Topic B	15.5	55.6	28.9
Topic C	19.0	54.9	26.1
Recall and Comprehension (all topics)	20.4	56.3	23.2
Recall and Comprehension (Topic A)	22.5	62.7	14.8
Recall and Comprehension (Topic B)	15.5	57.7	26.8
Recall and Comprehension (Topic C)	21.1	59.9	19.0
Value Concepts and Valuing Skills (all topics)	12.7	66.2	21.1
Inquiry Skills I (all topics)	21.1	54.2	24.6
Inquiry Skills II (all topics)	12.0	60.6	27.5

In examining the test results, the reader must keep in mind that a test score does not indicate why a particular performance occurred, but only that it did occur. After studying the results, the identification of reasons for that performance should be undertaken. There are a variety of factors that should be examined:

1. Student motivation. Consideration should be given to the degree to which students were motivated to perform to their levels of ability.
2. Student ability. While the statistical test of significance is designed to take into consideration fluctuations in the average ability levels of students, it is possible that a group of students with a particularly high or low average ability may come through a system. This is much more likely to be a factor in small systems than in large ones.
3. Readability. The achievement test was designed for a Grade 3 reading level. Jurisdictions should consider the average reading level of their Grade 3 students, as reading levels below Grade 3 will have an effect on test results that will be independent of achievement in social studies.
4. Teaching and curriculum. Consideration should be given to the type of instruction students have received in the jurisdiction and the adequacy of curricular implementation.

There will be other factors that are of importance in particular jurisdictions. School boards wishing to examine further the results in light of local factors are encouraged to establish their own local interpretation panels.

#### Absentee Rates

If more than 10% of the eligible students in a jurisdiction did not write the test, the reported averages for that jurisdiction may not accurately represent the true averages. Teacher-assigned marks for students who did not write could be compared with teacher-assigned marks for students who did write. If the averages are the same for the two groups, the reported achievement averages are probably representative. If the averages are different, some estimates can be made of what the achievement averages might have been if all students had written the test. Jurisdictions with high absentee rates may wish to contact the Student Evaluation Branch for assistance in estimating their averages.



## DATE DUE SLIP

DUE EDUC APR 17 '86

APR 08 RETURN

DUE EDUC SEP 18 '87

OCT 14 RETURN

DUE EDUC MAY 16 '91

MAY 14 RETURN

DUE EDUC AUG 06 '91

AUG 06 1991 RETURN

F255

JUN 12 1985

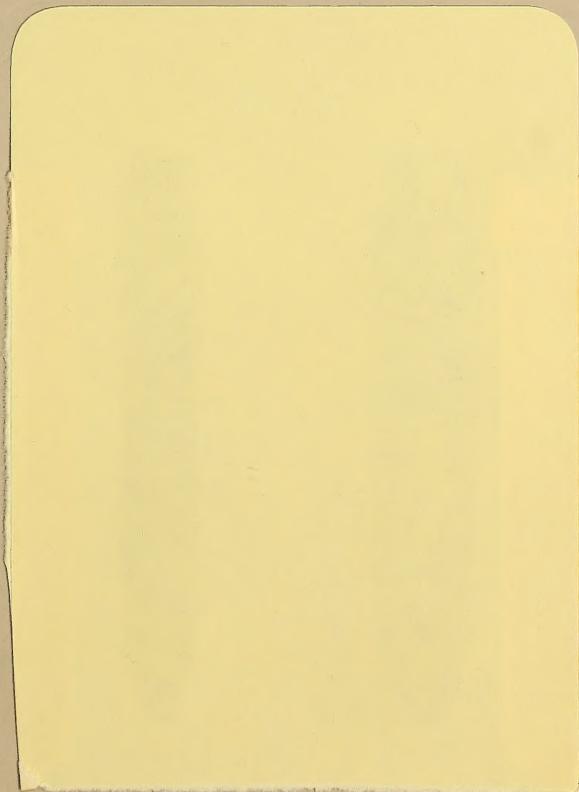
OCT 20 1987

LB 3054 C2 D3334 1984  
PROVINCIAL REPORT ACHIEVEMENT  
TEST GRADE 3 --

SERIAL M1 39904054 EDUC



\*000034691287\*



B4931 3